DOCUMENT RESUME

ED 272 876 CS 209 929

AUTHOR Kagel, Richard; Billings, LeRoy D.
TITLE Do Finished Television Commercials Test More Effectively Than Slide Show Storyboard Television Commercials: An Experimental Copy Test.
PUB DATE Aug 86
NOTE 12p.; Paper presented at the Annual Meeting of the Association for Education in Journalism and Mass Communication (69th, Norman, OK, August 3-6, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143) -- Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Advertising; Evaluation; *Mass Media Effects; *Media Research; Merchandising; *Television Commercials
IDENTIFIERS *Advertising Effectiveness

ABSTRACT

A study was conducted to investigate the theory that the more finished a television commercial is when tested, the more soundly it can be evaluated. In an experimental design for a health maintenance organization, four commercials were tested in both storyboard slide show and finished commercial formats in a controlled laboratory environment. Results revealed no significant differences in the effectiveness of the two formats. This finding indicates that advertisers do not benefit their campaign evaluation by undergoing the expense of producing the commercials for testing. (DF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED272876

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Do Finished Television Commercials Test More Effectively Than Slide Show Storyboard Television Commercials: An Experimental Copy Test

by

Richard Kagel, Ph.D.,
Associate Professor
and
LeRoy D. Billings,
Masters Candidate

Ted Kendall,
Research Assistant

Brigham Young University
Department of Communication
E-509 Harris Fine Arts Center
Provo, Utah 84602
801/378-2125

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Richard Kagel

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS209929

Do Finished Television Commercials Test More Effectively Than Slide Show Storyboard Television Commercials: An Experimental Copy Test

Abstract

In television commercial testing, advertisers have a constant concern that the degree to which their spots are finished will have a major influence on how well they are received by consumers. It is assumed that the more finished a commercial is when tested the more soundly it can be evaluated. In reviewing the literature on advertisement commercial copy testing, however, there is little to be found that substantiates this assumption. In a experimental design for a health maintenance organization which tested the same four commercials in both storyboard slide show and finished commercial formats in a controlled laboratory environment, the results showed that there was no significant difference between the two formats. The results of this test clearly indicate that advertisers do not significantly benefit their campaign evaluation by undergoing the expense of producing the commercials for testing. If these data hold up in a planned replication of this study using a different product category, behavioral intention measures, larger cell sizes and some other methodological refinements, then this result could have significant impact on the economics of television advertisement testing. Testing preproduced concepts has many obvious benefits, not the least of which would be to minimize ever skyrocketing costs of production; whereas, testing finished commercials, when there is no absolute performance difference, has no real benefit to either advertising agencies or to their clients.

Do Finished Television Commercials Test More Effectively Than Slide Show Storyboard Television Commercials: An Experimental Copy Test

Description and Hypotheses

In television commercial testing, advertisers have a constant concern that the degree to which their spots are finished will have a major influence on how well they are received by consumers. Positioning Advertising Copy Testing (PACT), a cumulation of dialogue among twenty-one leading American advertising agencies, reported in the Journal of Advertising Research:

> A good copy testing system recognizes that the more finished a piece of copy is, the more soundly it can be evaluated and requires, as a minimum, that alternative executions be tested in the same degree of finish.
>
> Experience has shown that the test results can often vary depending on the degree of finish of the test executions. Thus, careful judgment should be used in considering the importance of what may be lost in a less than finished version. Sometimes this loss may be inconsequential; sometimes it may be critical.
>
> The judgment of the creators of the advertising should be given great weight as to the degree of finish required to represent the finished advertisement for test purposes. If there is reason to believe that alternative executions would be unequally penalized in pre-production form, then it is generally advisable to test them in a more finished form.[1]

[1]Inserted booklet center section, The Positioning Advertising Copy Testing (PACT), p.20, Journal of Advertising Research, January/February, 1982.

It is assumed then that the more finished a commercial is when tested the more "soundly it can be evaluated." In reviewing the literature on advertisement commercial copy testing, however, there is little to be found that substantiates this assumption, perhaps because it is very expensive to do this kind of methodological research.

FHP, a health maintenance organization (HMO) based in California and Utah, was faced recently with the decision of moving ahead with a creative, preproduction proposal from their advertising agency to finance finished, ready-to-air commercials. The company was necessarily concerned about the associated high risk. These new commercials were meant to elevate the organization from advertising that established a high level of name recognition to a strategy that would move the consumer to positive action toward FHP. Specifically, the campaign goals were to:

1. Influence member company employees to investigate FHP
2. Influence member company employees to ask about FHP
3. Influence member employees to select FHP [2]

In order to minimize risk, FHP experimentally tested four commercials in a storyboard slide show format with a finished audio track to determine which of the four would yield the best behavioral intention response on the three stated criteria, using a six point Likert-type agreement scale. Based on the experimental results, no significance was found; and, generally, the commercials did no better than the control group which received no FHP commercial exposure. In addition, the mean scores showed little behavioral commitment toward FHP, the major objective of the campaign. FHP's

[2]From discussion with FHP marketing management.

advertising agency felt that the reason the commercials tested so poorly was the degree of finish used for the test, storyboard slide show versus finished production, and pressed to complete the commercials.

FHP consented to the production of all four commercials with the stipulation that they be post-tested in exactly the same fashion as the preproduction commercials. The finished test pool, counterpart to the storyboard test pool, was tested under the same criteria, and again, no significance was found. While this was an expensive decision for FHP and the results were certainly distressing to the advertising agency, it helped both to determine how well these commercials would contribute to the overall strategy. This would save money and disappointment in the long run. As a side benefit, it would allow the authors to test the proposition that the degree of finish of television commercials is important to the test outcome. Both sets of data, storyboard and finished, could be used in a factorial analysis of variance model to determine the main effect of the degree of finish factor. If there were no difference in the degree of finish, then advertisers could save themselves a great deal of effort, time and money by testing only preproduced concepts, a major benefit to both clients and their agencies. If there were a significant difference between the two formats, storyboard and produced commercials, and that difference were to weigh on the side of the finished commercials, then advertisers would need to consider the extra cost of testing a higher degree of finish as an important business and creative decision.

With this in mind, the research hypothesis (H_1) for this analysis is: Do finished FHP television commercials perform significantly better in an experimental setting using a factorial ANOVA model on behavioral intention scales directed toward the company, than do the same commercials in a

storyboard slide show format? The null hypothesis (H_0) is: There is no difference in the storyboard versus finished commercial formats based on the same criteria.

Methodology

In late 1984, Five experimental groups were randomly selected by a professional field work research company to view four FHP storyboard slide show commercial concepts. The fifth group, as control, saw no FHP spot. The respondents were screened and selected on broad demographic and employment criteria, descriptive of the company's target market. They were also screened for the usual involvement as recent research participants and affiliation with the media and advertising professions. The respondents were asked to complete a variety of tasks necessary to accomplish the experimental guise. This data also provided additional attitudinal comparisons to round out the analysis. These findings will not be reported in this presentation.

The second phase of the analysis, testing the produced commercials, took place in early 1985. This was accomplished using the same methodological criteria as the slide show test. The participant selection procedure, presentation and timing of treatments (all tightly scripted), data collection instruments, ordering effects and other extraneous variables were controlled for as rigorously as conditions allowed. All sessions were conducted by a professional research firm and moderator. The three dependent variables, investigate FHP, ask about FHP and select FHP were measured after the commercial exposure on a six point Likert-type agreement scale. Data tabulation and entry was carefully monitored and

checked. The analysis was accomplished using a Compaq Plus portable computer and the CRISP statistical package. The three dependent variables were all tested in three independent ANOVA procedures. The uneven cell sizes for each of the 10 groups averaged12 respondents.

The Findings

The following discusses the findings for each of the three dependent variables:

For the dependent variable: I plan to investigate FHP's Health Maintenance Organization (HMO): _____

Strongly Agree * (1)	Moderately Agree (2)	Slightly Agree (3)	Slightly Disagree (4)	Moderately Disagree (5)	Slightly Disagree (6)

*Numbering for tabulation purposes, not on original instrument.

The following is the ANOVA table resulting from the slide show and finish data:

Table 1

Summary of ANOVA Results for Independent Variable: Investigate

Source	df	SS	MS	F	P
Between Subjects	115	293.800			
Advertisements	4	15.145	3.800	1.46	0.2171 (ns)
Pre and finished	1	0.301	0.301	0.12	0.7337 (ns)
Interaction (AxP)	4	3.981	0.995	0.38	0.8206 (ns)**
Subjects w/Groups	106	274.331	2.588		

Collapsed mean for slide show (pre)= 3.464 (n=56)
Collasped mean for finished=3.333 (n=60) **see note

For the finish factor there is no significance, indicating that there was no real difference between the slide show test and the produced commercial test on this variable. The small difference between the two mean scores confirms this fact.

For the dependent variable: I plan to ask my employer about FHP:

Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Slightly Disagree

The following is the ANOVA table resulting from the slide show and finish data:

Table 2
Summary of ANOVA Results for Independent Variable: Ask

Source	df	SS	MS	F	P
Between Subjects	115	280.862			
Advertisements	4	23.453	5.900	2.40	0.0515 (ns)
Pre and finished	1	0.155	0.155	0.06	0.8001 (ns)
Interaction (AxP)	4	1.840	0.460	0.19	0.9431 (ns)**
Subjects w/Groups	106	255.410	2.409		

Collasped mean for slide show (pre)=3.518 (n=56)
Collasped mean for finished=3.417 (n=60) **see note

For the finish factor there is no significance, indicating that there was no real difference between the slide show test and the produced commercial test on this variable. The small difference between the two mean scores confirms this fact.

For the dependent variable: I am going to select FHP as my health Plan:

Strongly Agree	Moderately Agree	Slightly Agree	Slightly Disagree	Moderately Disagree	Slightly Disagree

The following is the ANOVA table resulting from the slide show and finish data:

Table 3

Summary of ANOVA Results for Independent Variable: Select

Source	df	SS	MS	F	P
Between Subjects	115	234.301			
Advertisements	4	23.249	5.800	2.96	0.0229 *
Pre and finished	1	0.370	0.370	0.89	0.6650 (ns)
Interaction (AxP)	4	2.643	0.337	0.19	0.8537 (ns)**
Subjects w/Groups	106	208.038	1.962		

Collasped mean for slide show (pre)=4.143 (n=56)
Collasped mean for finished=4.017 (n=60)

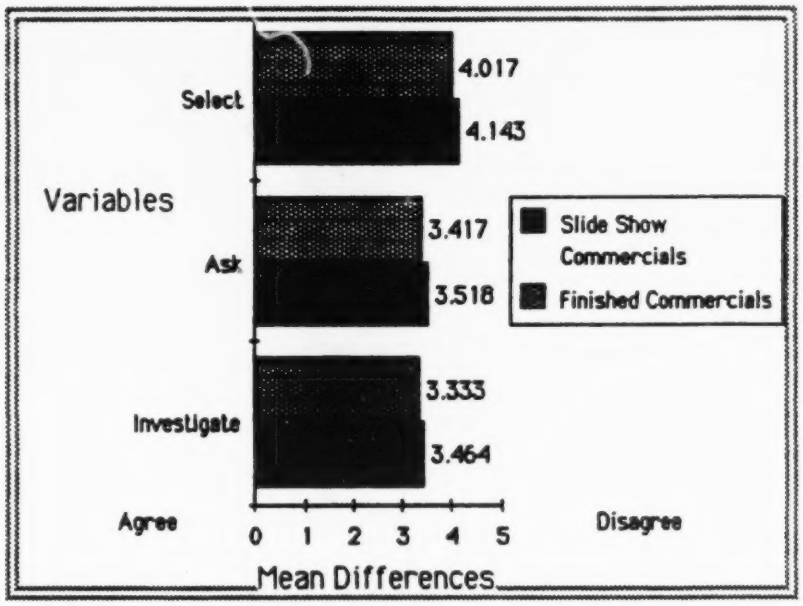
* The significance found on the advertisements factor indicate that when the preproduction (slide show) and finished commercials mean scores on this variable are combined there is a difference among groups, one that was not found when looking at the factor independently in the one-way ANOVA tests originally done. Through post hoc analysis, using the Scheffe and Tukey A and B tests, one commercial concept was found to perform significantly better than the others. This significance has no relationship to the test in question, the difference between slide show and finished commercials on the independent variable, the subject of this study.

**There was no significant interaction between the two factors of advertisements and slide show versus finished commercials, indicating no interaction to confound the results.

For the finish factor there is no significance, indicating that there was no real difference between the slide show test and the produced commercial test on this variable. The small difference between the two mean scores confirms this fact.

The results confirm that there is no significant difference between the slide show commercials and the finished commercials on any of the three dependent variables. The following chart graphically shows the small differences between mean scores:

Chart 1



Implications and Recommendations for Future Planned Research

The results of this test clearly indicate that there is no real difference in testing commercials in the preproduced or slide show stage from the finished or produced format. If these data hold up in a planned replication of this study using a different product category, behavioral intention measures, larger cell sizes and some other methodological refinements, then this result could have significant impact on the economics of television advertisement testing. Testing preproduced concepts has many

obvious benefits, not the least of which would be to minimize ever skyrocketing costs of production; whereas, testing finished commercials, when there is no absolute performance difference, has no real benefit to either advertising agencies or to their clients.